

Санкт-Петербургский государственный университет

Программная инженерия
Кафедра Системного программирования

Брыксин Матвей Александрович

Реконструкция трехмерной поверхности лица на базе морфированной модели

Бакалаврская работа

Научный руководитель:
к. ф.-м. н., доцент А. Т. Вахитов

Рецензент:
ген. дир. ООО “Биомоделирование” А. С. Петров

Санкт-Петербург
2016

SAINT-PETERSBURG STATE UNIVERSITY

Software Engineering
Software Engineering department

Matvey Bryksin

3D face reconstruction based on morphable model

Graduation Thesis

Scientific supervisor:
Associate professor Alexander Vahitov

Reviewer:
CEO "Biomodeling" Alexander Petrov

Saint-Petersburg
2016

Оглавление

Введение	4
1. Обзор	7
1.1. Нахождение перспективной камеры	7
1.2. Детекция ключевых точек на лице	8
1.3. Реконструкция поверхности лица	9
2. Описание подхода	11
2.1. Условия решаемой задачи	11
2.2. Начальное приближение камеры	11
2.3. Алгоритм Левенберга-Марквардта	12
2.4. Использование морфированной модели	13
3. Реализация	14
3.1. Система сбора статистических данных	14
3.1.1. Рендеринг трехмерной модели лица	16
3.1.2. Детектирование ключевых точек на лице	18
3.2. Система генерации тестов и проведения экспериментов .	19
3.3. Реализация предложенного подхода	21
4. Эксперименты	22
4.1. Зависимость от ковариационной матрицы	22
4.2. Задача нахождения перспективной камеры	24
4.3. Реконструкция поверхности лица и камеры	25
4.4. Зависимость от параметров тестовых данных	28
Заключение	29
Список литературы	30
Приложение	33

Введение

В современном мире графика в компьютерных играх и киноиндустрии продолжает совершенствоваться с каждым годом. Несколько последних лет прослеживается тенденция съемки настоящих актеров для отрисовки 3D моделей компьютерных персонажей. Для того, чтобы создать полнометражную компьютерную анимацию, требуется участие высоко квалифицированных художников, а также большое количество временных и вычислительных ресурсов. Поэтому развитие технологий захвата движений актера, построения трехмерных моделей сцены и объектов на ней, распознавания лиц и основных черт лица постепенно приводит к автоматизации процесса создания трехмерных моделей актеров и повышению их реалистичности.

Получение трехмерной модели лица обычно требует наличия специального устройства, как Kinect или лазерный сканер. Но такие приборы сложно использовать вне специальных лабораторий из-за трудностей в оборудовании необходимого окружения и их высокой цены. В более простом случае реконструировать 3D модель лица пытаются по одной или несколькими фотографиям. При таком подходе, как правило, используют базы шаблонных моделей для поиска максимально близких к исходному изображению ключевых областей.

Большинство ограничений в автоматических методах синтеза лиц существуют, чтобы избежать нереалистичных, непохожих на натуральное лицо, результатов. Например, некоторые методы требуют внесения ручных правок по ходу процесса генерации трехмерной модели или работают только для фронтального изображения лица. С появлением морфированной модели [1] (3D morphable model, 3DMM) стали активно развиваться новые направления автоматической генерации, реконструкции и анимации лиц. Основная идея использования морфированной модели заключается в разложении произвольной поверхности лица в линейную комбинацию базисных лиц. За счет аппроксимации произвольной модели базисными уменьшается требуемый размер заранее подготовленной базы, а высокая детализация исходных моделей

позволяет синтезировать реалистичные лица. Подход линейной аппроксимации не требует высокой вычислительной мощности, поэтому применяется даже для генерации моделей в реальном времени [8] и дает реалистичные результаты при построении анимации лиц.

Реконструкция трехмерной модели лица по двухмерному изображению состоит из двух задач: построения трехмерной поверхности лица и нахождения его положения в пространстве. В соответствии с определением параметрической модели Парке [18] поверхность лица задается набором координат трехмерных вершин и множеством плоскостей, проходящих через них. Положение трехмерной модели в пространстве описывается матрицей поворота и вектором переноса камеры.

Задача определения ориентации и позиции откалиброванной перспективной камеры по n известным 3D точкам в координатах модели и их 2D проекциям на изображении (perspective- n -point problem, PnP) имеет решение при $n \geq 3$ [4, 7]. При наличии информации о более, чем 3 точках на изображении и их положении в трехмерном пространстве, глобально оптимальное решение этой задачи может быть найдено алгоритмом OPnP [26] за вычислительную сложность равную $O(n)$.

Минимизация функции стоимости в виде нелинейных наименьших квадратов методом Левенберга-Марквардта позволяет уменьшить ошибку найденной ориентации и позиции камеры OPnP алгоритмом. Для этого в качестве весовой матрицы в методе наименьших квадратов следует взять обратную ковариационную матрицу распределения 2D точек в 3D. Алгоритм Левенберга-Марквардта помимо оптимизации параметров камеры можно в тоже время использовать для приближения параметров морфированной модели, с целью получения трехмерной поверхности лица.

Описанные положения о вычислении начального приближения параметров перспективной камеры с помощью OPnP алгоритма и использование обратной ковариационной матрицы в качестве матрицы весов при минимизации функции стоимости в виде нелинейных наименьших квадратов легли в основу предлагаемого статистического подхода.

Постановка задачи

Цель данной работы реализовать и провести апробацию нового статистического подхода по реконструкции трехмерной поверхности лица и нахождению матрицы поворота и вектора переноса перспективной камеры по одному исходному изображению повернутого лица.

В ходе работы для достижения описанной цели были поставлены следующие задачи.

1. Разработать систему сбора статистических данных, необходимых для предлагаемого алгоритма.
2. Реализовать предлагаемый подход к реконструкции 3D поверхности лица и нахождению параметров перспективной камеры.
3. Разработать систему генерации тестовых данных и проведения сравнительных экспериментов.
4. Провести сравнение с существующими алгоритмами нахождения перспективной камеры и алгоритмами реконструкции 3D поверхности лица.
5. Исследовать зависимость результатов предлагаемого подхода от изменения параметров тестовых данных.

1. Обзор

Предметная область реконструкции трехмерных моделей имеет особенность, что большинство алгоритмов узко специализированны на решение задачи под определенные исходные данные или используют собственную базу шаблонов трехмерных лиц, которая не является общедоступной. Поэтому основная задача обзора заключалась в исследовании условий использования алгоритмов и анализа возможности их применения для задачи в поставленном виде. Далее в каждой секции представлен обзор существующих алгоритмов.

1.1. Нахождение перспективной камеры

Большинство алгоритмов для нахождения ориентации и положения камеры основаны на соответствии точек или линий. Самый стандартный подход к калибровке камеры формулируется для 2D-3D соответствия точек [7]. Для решения этой задачи необходимы данные как минимум о 3 точках [4]. Сначала алгоритмы решали задачу для фиксированного числа точек от 3 до 5, но с увеличением их количества возрастала точность восстановления камеры. Первым $O(n)$ алгоритмом, решающим задачу для 5 и более точек, был EPnP [13]. Основная идея заключалась в том, чтобы представить 3D координаты как линейную комбинацию 4 контрольных точек, которые принимались за неизвестные независимо от общего количества исходных точек. Эти контрольные точки были получены при помощи простых техник линеаризации. Линеаризация была последовательно заменена полиномиальными решениями в Robust PnP [14] и далее в алгоритме OPnP (Optimal Pose from n Points) [26], который дает в общем случае наиболее точные результаты. В OPnP для поворота используется параметризация кватернионами, что позволяет оценивать позицию камеры при помощи решения базиса Грёбнера.

Итеративные подходы формулируют нахождение параметров камеры как задачу минимизации в форме нелинейных наименьших квадратов и минимизируют геометрическую или алгебраическую ошибку.

Один из первых таких методов был алгоритм ортогональной итерации (Orthogonal Iteration Algorithm, OI) [16]. Он, в отличие от множества других алгоритмов, минимизирует ошибку, описанную в 3D координатах модели, а не пытается стандартно минимизировать 2D ошибку перспективного преобразования [6]. На практике алгоритм сходится быстро, но может остановиться в неподходящем локальном минимуме, если был некорректно инициализирован. Основным недостатком алгоритма ортогональной итерации в том, что из-за итеративности он работает намного медленнее аналогов и показывает менее стабильные результаты.

1.2. Детекция ключевых точек на лице

Большинство существующих подходов по реконструкции лиц базируются на алгоритмах нахождения ключевых точек на лице (Facial feature point detection, FFPD). Поэтому необходимо было найти наиболее точный алгоритм из существующих подходов к детектированию особых точек на лице, подходящий под условия поставленной задачи.

Алгоритмы детектирования ключевых точек на лице относятся к классу методов обучения с учителем или частичного обучения на вручную размеченных изображениях. Согласно классификации, предложенной в работе Cootes [2], ключевые точки на лице могут быть разделены на 3 типа: точки, устойчивые к повороту, такие как центр глаза, края глаз или губ; точки, неустойчивые к повороту, которые рассчитываются как экстремум на кривой соединяющей устойчивые точки; и точки, интерполированные из 2 предыдущих типов - по краям лица и на подбородке. Последний тип в силу интерполяции сохраняет меньше всего информации о повороте и положении лица.

Основной подход к нахождению ключевых точек на лице заключается в локализации модели каркаса относительно появления лица на изображении. Как правило, нахождение ключевых точек состоит из 2 фаз: тренировки, когда происходит обучение на позициях точек и подстраивание модели каркаса под них, и тестирования, когда обученная модель применяется к исходному изображению, чтобы найти положе-

ния ключевых точек. Обычно процесс поиска каркаса начинается с грубого начального значения всего изображения целиком, но в некоторых методах используется поиск положения лица [15, 22], который обеспечивает более близкое начальное значение.

В множестве существующих подходов выделяется 4 направления методов детектирования ключевых точек: Constrained Local Model (CLM), Active Appearance Model (AAM), регрессионные методы и другие.

CLM-методы рассматривают области появления каждой ключевой точки независимо от других [21]. Для этого положение каждой точки внутри возможной области определяется соответствующим ей локальным экспертом. Финальное положение точки определяется после сопоставления полученного положения точек друг относительно друга с заранее обученной моделью каркаса. В отличие от них AAM-методы используют целостный подход для всех точек [24]. Модель каркаса точек и их области появления строятся как линейная комбинация базиса, обученного на тренировочных изображениях. Регрессионные методы определяют общее положение точек без обучения каких-либо моделей.

Детальный анализ существующих подходов [23] показывает, что каскадный набор простых линейных регрессоров позволяет достигать наиболее точных результатов [25, 5]. Такие подходы достаточно устойчивы к изменению положения точек относительно друг друга, что позволяет им обеспечивать стабильность на повернутых изображениях лиц.

1.3. Реконструкция поверхности лица

Для реконструкции поверхности лица популярны алгоритмы, использующие Kinect, лазерный сканер или другие устройства, сохраняющие информацию с учетом глубины [9]. Но большинство существующих подходов пытаются решать задачу без специальных устройств и подготовленного помещения для одного или множества исходных изображений с неоткалиброванной камеры [11]. Среди таких алгоритмов можно выделить два направления: методы восстановления поверхности только по исходным изображениям и методы, использующие базы

трехмерных лиц для построения полной модели лица.

Получение поверхности из освещения (Shape from shading, SFS) — основное направление в реконструкции трехмерной поверхности без использования базы трехмерных моделей лиц [12]. Идея алгоритма заключается в использовании теней на изображении для вычисления интенсивности и расчета направления нормалей поверхностей для освещения Ламберта. Но такие подходы сильно зависят от света на изображении и получают поверхность только видимой части лица.

С появлением морфированной модели [1] процесс генерации трехмерных моделей стал более простым и быстрым, что способствовало развитию алгоритмов в направлении синтеза лиц [19]. Реконструкция трехмерной поверхности лица с использованием морфированной модели сводится к подбору параметров, описывающих ее поверхность. Большинство таких алгоритмов итеративные и базируются на ключевых точках на лице, а положение лица в пространстве находят заранее. Один из таких алгоритмов [20] базируется на минимизации ошибки устойчивых ключевых точек и размеченных адаптивных контурах положения неустойчивых точек по краям лица. Однако, этот алгоритм в оригинальной статье был реализован только для ортографической камеры. Метод реконструкции трехмерной поверхности и текстуры лица, предложенный в работе D.Jiang [10], показывает результаты высокой точности на реальных изображениях. Этот алгоритм имеет узкое применение, так как требует фронтального изображения лица и неустойчив к повороту. В отличие от аналогов, которые минимизируют ошибку проектирования 3D точек, в алгоритме P.Huber [8] подбирает параметры морфированной модели на основе ключевых точек на изображении. Такой подход позволяет использовать информацию о положении лица, полученную из каскадного детектора ключевых точек.

В обзоре рассмотрены существующие подходы к реконструкции трехмерной поверхности лица и нахождению параметров камеры. Для сравнения с предлагаемым подходом был выбран ОРnР алгоритм, алгоритм ортогональной итерации и базовый алгоритм ошибки перспективного преобразования.

2. Описание подхода

2.1. Условия решаемой задачи

Реконструкция лица для перспективной камеры заключается в решении двух задач компьютерного зрения. Первая — определение положения и ориентации лица, что равносильно нахождению матрицы поворота и вектора переноса камеры. Задача формулируется для набора двумерных точек на изображении и соответствующих им трехмерным координатам. Вторая — реконструкция трехмерной поверхности лица. Для этого необходимо найти трехмерные координаты точек модели и множество плоскостей проходящих через них.

Статистический подход применяется как для одной задачи нахождения параметров камеры, так и для первой и второй задачи одновременно. При реконструкции трехмерной поверхности и нахождении параметров камеры для получения трехмерной модели в предлагаемом подходе используется морфированная модель.

2.2. Начальное приближение камеры

Предложенный подход состоит из нескольких подготовительных действий, которые выполняются до запуска алгоритма. В первую очередь, находим ключевые точки на изображении и полученный набор нормализуем при помощи обратной матрицы калибровки камеры. Сопоставляем ключевым точкам соответствующие им трехмерные координаты по индексам в средней морфированной модели. Индексы получены заранее по не повернутому изображению средней модели для используемого алгоритма детекции особых точек.

После этого для набора двумерных точек и их трехмерных координат применяем алгоритм OPnP и получаем начальное приближение матрицы поворота R_0 и вектора переноса t_0 камеры. Это начальное приближение используем в качестве начального значения в алгоритме Левенберга-Марквардта.

2.3. Алгоритм Левенберга-Марквардта

В общем случае для приближения параметров камеры и трехмерной поверхности лица можно сформулировать задачу минимизации суммы квадратов остатков (1), где e — вектор остатков, W — симметричная положительно определенная весовая матрица.

$$\sum e^T W e \rightarrow \min \quad (1)$$

В предлагаемом подходе за вектор остатков принимается ортогональная ошибка, составляющая основную идею итеративного алгоритма ортогональной итерации [16]. Согласно нему вектор остатков рассчитывается как $e = M X_{cam}$, где $X_{cam} = R X(\alpha) + t$ — трехмерные точки в координатах морфированной модели. Для этого по каждой сдетектированной точке в однородных координатах $x_i = (u \ v \ 1)^T$ рассчитаем оператор (2), где $\bar{x}_i = \frac{x_i}{\|x_i\|}$.

$$L_i = \bar{x}_i \bar{x}_i^T \quad (2)$$

Строим блочно-диагональную матрицу M для расчета трехмерной ошибки из множества лежащих на диагонали L_i , рассчитанных для каждой ключевой точки.

$$M = \begin{pmatrix} I - L_1 & 0 & \cdots & 0 \\ 0 & I - L_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I - L_n \end{pmatrix} \quad (3)$$

Основная идея предлагаемого подхода заключается в использовании обратной ковариационной матрицы в качестве весовой матрицы W в алгоритме Левенберга-Марквардта [17]. Известно, что если в качестве весовой матрицы W используется обратная ковариационная матрица V случайных ошибок (то есть $W = V^{-1}$), то оценки параметров, найденные с помощью метода наименьших квадратов (МНК), будут наиболее эффективными. Ковариационная матрица описывает распределение ключевых точек в координатах модели.

Полученное начальное приближение матрицы поворота R_0 используется для перевода матрицы ковариации из координат модели в координаты камеры. Перевод матрицы в координаты камеры рассчитывается покомпонентно (4) для каждой сдетектированной точки $i, j \in 1..n$.

$$\Sigma_{ij}^{cam} = R_0 \Sigma_{ij} R_0^T \quad (4)$$

В нашем случае ковариационная матрица в координатах камеры $V = M \Sigma^{cam} M^T$ является вырожденной (сингулярной или особой), поэтому в алгоритме вместо обратной матрицы используется псевдообратная матрица V^+ , которая рассчитывается при помощи SVD разложения.

2.4. Использование морфированной модели

Морфированная модель — это заданная параметрически трехмерная модель, представленная линейной комбинацией трехмерных поверхностей лиц из подготовленной базы. Она математически описывается средней моделью лица $\mu \in R^{3m}$ и вектором случайных величин $\alpha_i \in N(0, \sigma_i^2)$ с известной дисперсией $\sigma \in R^{n-1}$ для соответствующих ортонормированных базисов $U = [u_1, \dots, u_n] \in R^{3m \times n-1}$, описывающих компоненты морфированной модели (5).

$$shape = \mu + U diag(\sigma) \alpha \quad (5)$$

В нашем подходе ключевым точкам сопоставляются координаты в средней модели μ для поиска начального решения OPnP. В случае реконструкции трехмерной поверхности средняя модель принимается за начальное значение алгоритма. На каждой итерации генерируется трехмерная модель по найденным значениям α_i и подбираются новые значения случайных величин.

Полученные алгоритмом значения компонент морфированной модели применяются для генерации трехмерной модели лица, а найденные матрица поворота и вектор переноса для отрисовки лица в восстановленном положении.

3. Реализация

Для реализации предлагаемого подхода на языке C++ с использованием библиотек OpenGL и OpenCV была разработана система сбора статистических данных, необходимых для расчета ковариационной матрицы, лежащей в основе предлагаемого подхода. Подход вместе со системой проведения экспериментов разработаны на языке Matlab, аналогично алгоритмам, с которыми проводилось сравнение. Для проведения экспериментов на основе модулей, разработанных в рамках системы сбора статистики, была создана система генерации тестовых данных. С целью проведения сравнения предлагаемого подхода с аналогами был реализован алгоритм ортогональной итерации [16] и базовый алгоритм подсчета ошибки перспективного преобразования [6].

3.1. Система сбора статистических данных

Основная цель создания системы сбора статистики заключалась в получении данных о распределении ключевых точек в трехмерном пространстве в координатах морфированной модели. Данные нужны для расчета ковариационной матрицы, лежащей в основе предлагаемого подхода. Цикл получения статистики по ключевым точкам на изображении одной модели состоит из прохождения нескольких последовательных шагов, представленных на рисунке 1.

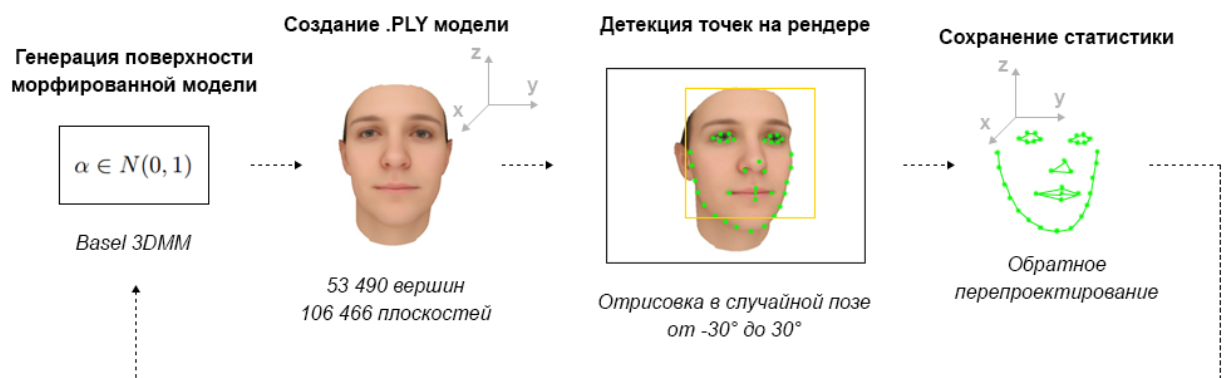


Рис. 1: Модули системы сбора статистики

Генерация трехмерной поверхности лица происходит при помощи морфированной модели из инструментария 3D Basel Face Model (BFM) на языке Matlab. Для этого создается вектор α из 199 компонент, где $\alpha_1 \in N(0, 1)$ является случайной величиной с нормальным распределением, а все остальные значения нули. Первый коэффициент отвечает за форму лица, поэтому его изменение позволяет генерировать одинаковые лица с разной формой. Полученный вектор используется для генерации лица стандартной функцией инструментария BFM.

Полученная трехмерная модель лица сохраняется в файл в соответствии с .PLY форматом описания трехмерных моделей. По каждой из 53490 вершин модели сохраняется информация о ее координате и цвете, после чего записывается информация о 106466 плоскостях, проходящих через эти вершины.

Сохраненная трехмерная модель подается на вход рендеру, который отрисовывает лицо на сцене при помощи примитивов OpenGL в случайном положении с поворотом лица в диапазоне от -30° до 30° . Далее на полученном изображении лица детектируются ключевые точки. Подробная реализация рендеринга и нахождения ключевых точек описаны далее в следующих секциях.

Для сбора статистики распределения ключевых точек в пространстве необходимо сопоставить 2D ключевой точке 3D координату вершины в модели. По известным x и y координатам на изображении сцены восстанавливается глубина точки в z -буфере при помощи функции `glReadPixels`. Далее при помощи Modelview, Projection, Viewport матриц и стандартной функции `gluUnProject` происходит обратное проектирование 2D точки в 3D координаты модели. В случае, когда найденная ключевая точка не попала на лицо, 3D координату восстановить невозможно, поэтому был добавлен пропуск таких исключений.

В полученных обратно данных 3D точек существует погрешность, которая объясняется недостаточно высоким разрешением изображения, ошибкой сохранения данных в 32 битном z -буфере или просто потерей точности при перемножении матриц. Поэтому найденная 3D точка не совпадает с координатами вершин на модели. С целью исправить по-

грешность был реализован поиск ближайшей вершины, с минимальным расстоянием от найденной. Для проверки корректности работы был создан тест, в котором найденная вершина модели снова отрисовывается на изображении. В среднем ошибка отрисованной вершины составляет 1.23 пикселя на разрешении изображений 800 на 600.

В результате сохраняется статистика о положении ключевых точек в трехмерном пространстве. Информация о найденных 3D координатах для каждой ключевой точки сохраняется в файл по каждой модели. Эти действия повторяются в цикле на множестве моделей. Была собрана статистика по 3000 моделей с малой, средней и большой вариацией первого коэффициента α_1 .

Для предлагаемого подхода был реализован расчет ковариационной матрицы по собранной статистике. Матрица рассчитывается покомпонентно для каждой сдетектированной точки по формуле (6), где X_i, X_j - трехмерные координаты модели, соответствующие двумерным проекциям.

$$\Sigma_{ij} = E[(X_i - EX_i)(X_j - EX_j)^T] \quad (6)$$

В результате работы системы сбора статистики было рассчитано 4 матрицы ковариации по 1000 моделей с малой вариацией первого коэффициента, по 1000 со средней, по 1000 с большой и по 1000 с разной вариацией от малой до большой. Матрицы были рассчитаны для сравнения результатов предлагаемого подхода от вариации статистических данных.

3.1.1. Рендеринг трехмерной модели лица

Модуль рендеринга трехмерного лица состоит из нескольких последовательных этапов: загрузка объектов на сцену, инициализация параметров отрисовки, инициализация и создание окна сцены, отрисовка объектов на сцене, сохранение изображения сцены, очистка графического буфера и буферов объектов на сцене. Также рендер имеет набор методов, которые позволяют получать параметры созданного изоб-

ражения, такие как фокусное расстояние камеры, матрица поворота, вектор переноса и другие.

Трехмерная модель загружается из .PLY файла при помощи библиотеки Assimp, которая автоматически рассчитывает для каждой вершины направление ее нормали, что необходимо для реалистичной отрисовки модели без текстуры. Далее происходит инициализация параметров отрисовки: отрисовка с текстурой или без (пример на рисунке 2), разрешение окна, угол поворота и смещение модели, уровень Гауссовского шума на изображении.

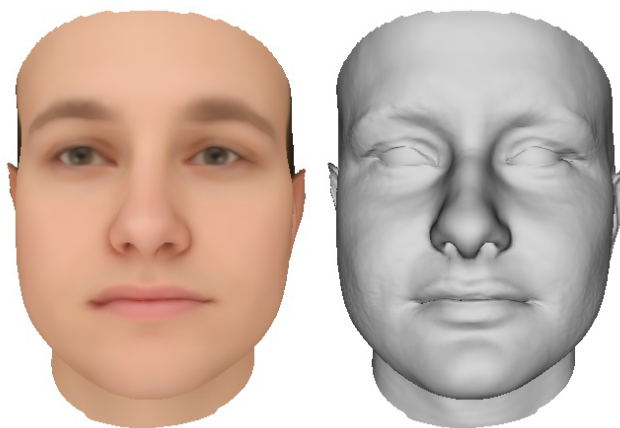


Рис. 2: Отрисовка лица с текстурой и без текстуры

Самый стандартный подход к отрисовке сцены при помощи библиотеки OpenGL это использовать для отрисовки бесконечный цикл `glutMainLoop()`. Однако, этот способ не позволяет изменять параметры созданного окна сцены, что критично для проведения эксперимента зависимости точности подхода от разрешения изображений. Более того, такой способ постоянно перерисовывает сцену и является отдельным процессом в системе, а в нашем случае требуется только однократно отрисовать каждое лицо. Для решения проблемы используется искусственный вызов события отрисовки `glutMainLoopEvent`, который поддерживается в библиотеке OpenGL только для языка C++. После события отрисовки и сохранения изображения сцены окно сцены может быть успешно закрыто, чего нельзя сделать в бесконечном цикле. Этот способ позволяет однократно проинициализировать рендер функцией

glutInit и многократно полностью менять параметры отрисовки.

Отрисовка трехмерной модели происходит при помощи примитива GL_TRIANGLES, что на порядок медленнее, чем при использовании вершинных и фрагментных шейдеров, но он проще в реализации, а в нашем случае для сбора статистики не требуется скорость, так как расчет проводится заранее. Поэтому для каждой вершины передается ее координата и цвет, а в случае отсутствия текстуры еще и нормаль в точке. При отрисовке без текстуры для создания освещенности на сцене используется стандартное освещение GL_LIGHT0.

3.1.2. Детектирование ключевых точек на лице

Для нахождения ключевых точек на лице был выбран SDM алгоритм каскадной регрессии [25], который по сравнению с другими дает наиболее стабильные результаты при повороте. Алгоритм имеет готовую реализацию в открытом доступе на языке C++ и применяется для детекции 68 ключевых точек. Из этого набора были выбраны 22 ключевые точки, которые наиболее устойчивы к повороту лица, по краям глаз, губ и на носу. Пример найденных ключевых точек на лице представлен на рисунке 3.

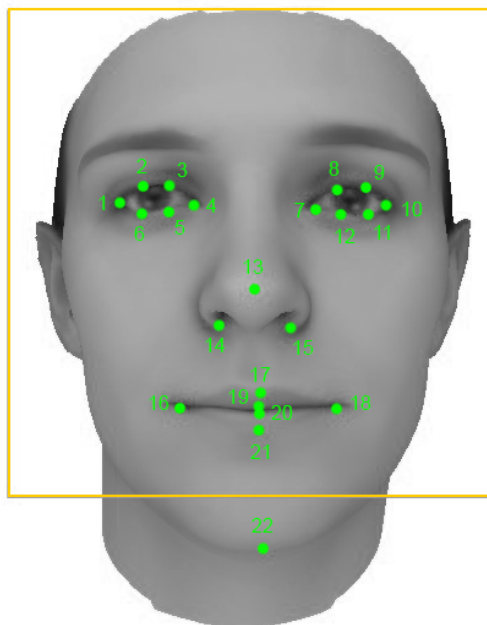


Рис. 3: Ключевые точки, устойчивые к повороту лица

В связи с предположением использования только устойчивых ключевых точек в SDM алгоритм была добавлена оптимизация поиска лица Нааг классификатором [15], который дает более близкое начальное значение для каскадной регрессии. Для этого использовался классификатор `haarcascade_frontalface_alt.xml` из библиотеки OpenCV и встроенная функция `detectMultiScale`.

С учетом добавленной оптимизации алгоритм был обучен для дескриптора особых точек HOG (Histogram of Oriented Gradients) [3] на открытой базе изображений и ключевых точек IBUG из отобранных 20 изображений, содержащих только одно лицо. HOG дескриптор вычисляется для 4 квадратных ячеек вокруг каждой ключевой точки, где высота одной ячейки 12 пикселей. Гистограмма ориентации каждой ячейки разбивается на 4 интервала по 90° .

Детектор был обучен для 3 последовательных линейных регрессоров. Параметры каждого регрессора вычисляются как решение линейной системы $A*x = b$, где A - матрица тренировочных данных, x - обучаемые коэффициенты, найденное методом наименьших квадратов из открытой математической библиотеки Eigen для языка C++. Следует отметить, что коэффициенты для каждой ключевой точки обучаются независимо, а обучение следующего регрессора происходит после получения предсказаний уже обученного регрессора на уровень выше.

Подобранные после обучения коэффициенты линейной модели для 22 ключевых точек были сохранены в бинарный файл, который далее каждый раз загружался при инициализации детектора и использовался для предсказания положения ключевых точек на изображении лица.

3.2. Система генерации тестов и проведения экспериментов

Для проведения сравнительных экспериментов на основе разработанных модулей из системы сбора статистики была разработана система генерации тестов и проведения экспериментов. Для этого были описаны требования к тестовым данным, необходимые для сравнения алгорит-

мов друг с другом. Для проведения сравнения необходимо иметь информацию о фокусном расстоянии камеры, ключевых точках на изображении, разрешении и уровне зашумленности изображения, матрицу поворота и вектор переноса, а также информации о трехмерной поверхности лица. В силу сложности формирования базы тестовых трехмерных поверхностей реальных лиц, для сравнения результатов реконструкции использовались параметры лиц морфированной модели.

Для каждого теста для проведения экспериментов в файл сохраняется информация в соответствии со спецификацией:

- Фокусное расстояние камеры – float
- Количество ключевых точек – int
- Ключевые точки – $\langle \text{int}, \text{int} \rangle []$
- Разрешение изображения – $\text{int}[2]$
- Гауссовский шум – float
- Матрица поворота – $\text{float}[9]$
- Вектор переноса – $\text{float}[3]$
- Размерность морфированной модели - int
- Компоненты морфированной модели - $\text{float}[]$

Система проведения экспериментов на языке Matlab считывает информацию из файла и запускает тестирование набора алгоритмов на тестовых данных. В результате работы алгоритмов получаются значения R_{real} и t_{real} , а в случае реконструкции трехмерной поверхности еще и вектор компонент морфированной модели α_{real} . Для полученных результатов и истинных значений параметров из тестового файла рассчитывается ошибка поворота камеры в градусах (7), ошибка вектора

переноса в процентах (8) и суммарная ошибка компонент морфированной модели с учетом распределения (9), где $\alpha_i \in N(0, \sigma_i^2)$.

$$E_{rot} = \max_{k=1..3} \text{acos}(\text{dot}(R_{truth}^k, R_{real}^k)) * 180/\pi \quad (7)$$

$$E_{trans} = ||t_{truth} - t_{real}|| / ||t|| * 100 \quad (8)$$

$$E_{alpha} = \sum_{i=1}^n \sigma(i)^2 * (\alpha_{truth}(i) - \alpha_{real}(i))^2 \quad (9)$$

3.3. Реализация предложенного подхода

Статистический подход к реконструкции трехмерной поверхности лица был реализован на языке Matlab и использует готовую реализацию алгоритма Левенберга-Марквардта из функции `lsqnonlin` в базовом пакете Optimization toolbox. Работа алгоритма ограничивается 20 итерациями, с целью уменьшения времени работы. Подход реализован для задачи нахождения параметров перспективной камеры, а также для задачи реконструкции поверхности и камеры одновременно. С целью уменьшения параметров в алгоритме и оптимизации его работы матрица поворота из 9 значений параметризуется 3 углами при помощи функции `rodrigues`.

4. Эксперименты

Для сравнения результатов нахождения параметров перспективной камеры взят OPnP алгоритм [26], который среди аналогов показывает наиболее точные результаты. С целью проведения экспериментов по реконструкции поверхности лица и камеры был реализован итеративный алгоритм подсчета ортогональной ошибки [16] (на графиках OPnP+OI), который в отличие от предложенного подхода использует в качестве весовой матрицы единичную. Также был реализован базовый алгоритм подсчета ошибки перспективного преобразования [6] (на графиках OPnP+Proj), ошибка которого рассчитывается для $X_{proj} = K * (R * X(\alpha) - t)$ по формуле (10), где K - матрица калибровки камеры, а $X(\alpha)$ - вершина в морфированной модели.

$$e_{proj} = \begin{pmatrix} \frac{X_{proj}(1)}{X_{proj}(3)} - x \\ \frac{X_{proj}(2)}{X_{proj}(3)} - y \end{pmatrix} \quad (10)$$

Оба реализованных итеративных алгоритма за начальное значение принимают решение, найденное OPnP алгоритмом. Ошибка предлагаемого подхода на всех графиках выделена зеленым цветом.

4.1. Зависимость от ковариационной матрицы

В результате работы системы сбора статистики была получена информация о распределении ключевых точек в трехмерном пространстве по 3000 моделей со случайным поворотом от -30° до 30° . Было рассчитано 4 ковариационных матрицы, которые отличаются вариацией случайных величин в компонентах α морфированной модели, по которым собиралась статистика. Первая матрица была рассчитана по 1000 моделей с малой вариацией первого коэффициента (от -2 до 2), вторая по 1000 со средней (от -4 до -2 и от 2 до 4), третья по 1000 с большой (от -6 до -4 и от 4 до 6), а четвертая по 1000 моделям с разной вариацией от малой до большой. Исследования подхода для решения задачи нахождения камеры проводились на 100 моделях со случайным первым коэффициентом

$\alpha_1 \in N(0, 1)$ и случайным поворотом от -30° до 30° . Результаты средней ошибки предлагаемого подхода представлены на рисунке 4 ошибки матрицы поворота в градусах (формула 7) и вектора переноса в процентах (формула 8). Графики медианы ошибки представлены в Приложении на рисунке 14.

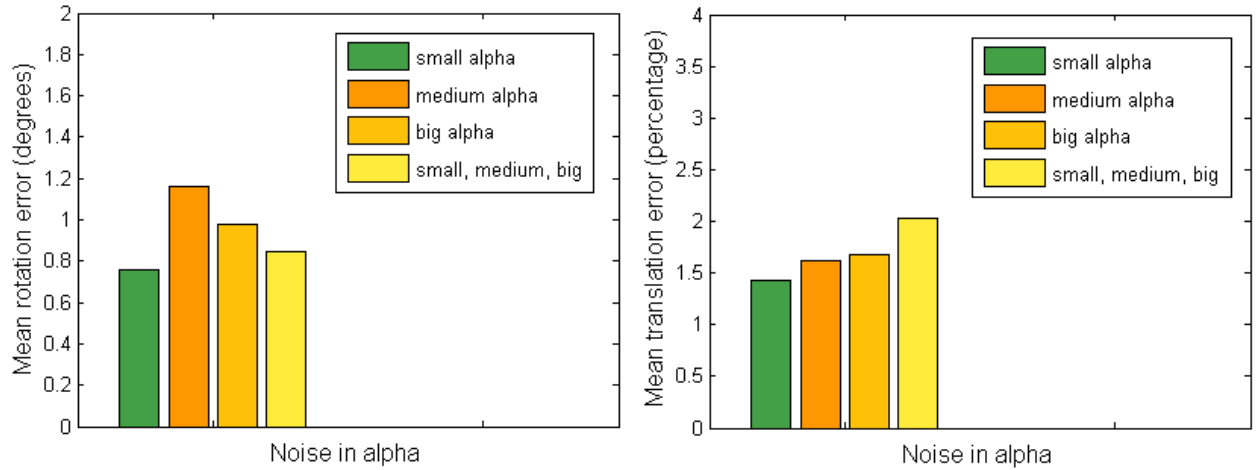


Рис. 4: Средняя ошибка камеры от разной матрицы ковариаций

Также была исследована зависимость от ковариационной матрицы результатов предлагаемого подхода для задачи реконструкции трехмерной поверхности и параметров камеры. Исследования проводились на 50 моделях со всеми случайными компонентами морфированной модели $\alpha_i \in N(0, 1), i = 1..199$ и случайным поворотом от -30° до 30° . Результаты средней ошибки реконструкции поверхности лица (формула 9) и камеры представлены на рисунке 5. Графики медианы ошибки представлены в Приложении на рисунке 15.

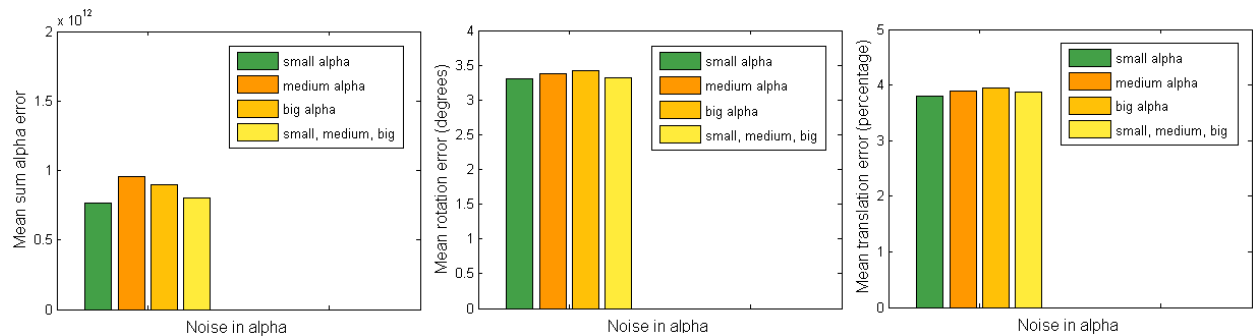


Рис. 5: Средняя ошибка реконструкции поверхности и камеры от разной матрицы ковариаций

Из рисунков 4 и 5 можно сделать общий вывод, что наиболее точные и стабильные результаты были получены с использованием матрицы ковариации с малой вариацией поверхности лица. Это может быть следствием того, что лица со средней и большой вариацией выглядят менее реалистично.

Сравнение предлагаемого подхода, представленное в следующих главах, проводилось с использованием ковариационной матрицы с малой вариацией.

4.2. Задача нахождения перспективной камеры

Для сравнения алгоритмов было проведено исследование зависимости результатов от угла поворота лица. Для этого взята одна модель со случайным первым коэффициентом морфированной модели $\alpha_1 \in N(0, 1)$ и остальными компонентами разными нулю. Модель равномерно поворачивается с шагом в 10° в диапазоне от -30° до 30° . Результаты для задачи нахождения перспективной камеры представлены на рисунке 6 ошибки матрицы поворота и вектора переноса.

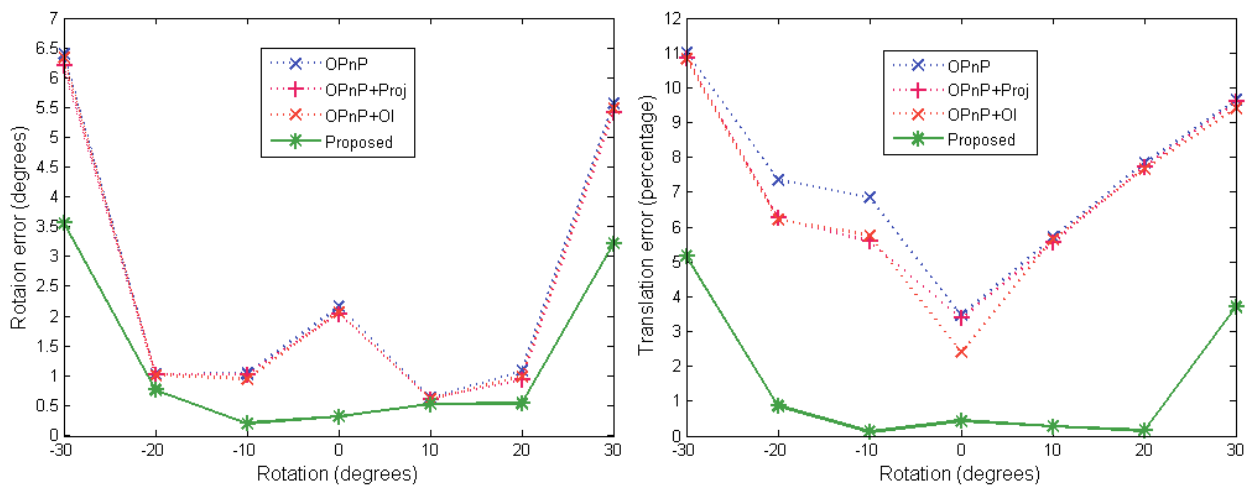


Рис. 6: Зависимость ошибки поворота и переноса от поворота лица

На графиках видно, что предлагаемый статистический подход показывает наиболее точные результаты. Следует особенно отметить стабильность результатов при повороте лица до 20° и устойчивость к сильному повороту на 30° .

Для подсчета средней ошибки и медианы было подготовлено 100 тестов на основе трехмерных лиц со случайной первой компонентой морфированной модели $\alpha_1 \in N(0, 1)$ на случайном повороте от -30° до 30° . Средняя ошибка матрицы поворота и вектора переноса представлена на рисунке 7. Графики медианы ошибки представлена в Приложении на рисунке 12.

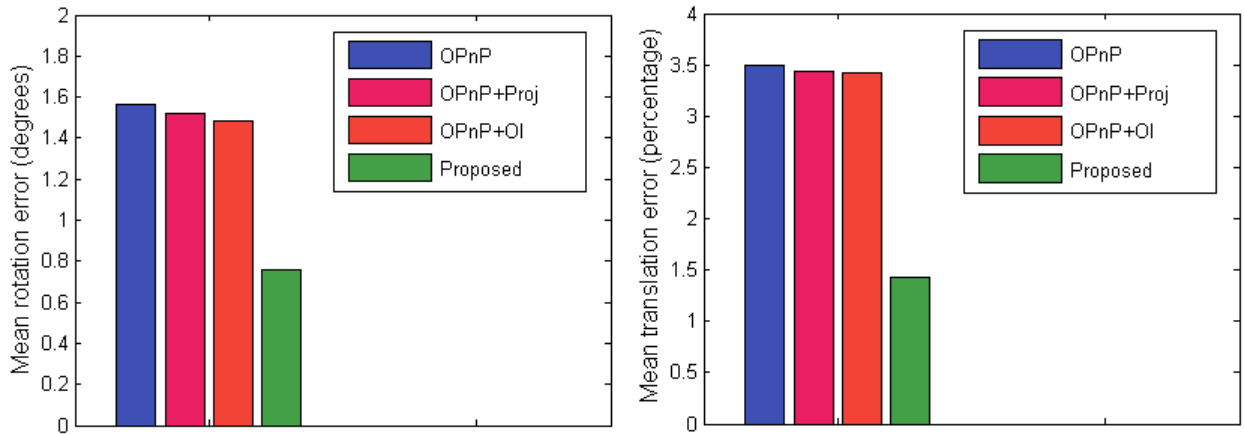


Рис. 7: Средняя ошибка поворота и переноса на 100 тестов

На рисунке видно, что ошибка предлагаемого подхода для определения поворота меньше 1° и составляет около 1.5% для переноса. Полученный результат почти в 2 раза точнее всех алгоритмов, с которыми проводилось сравнение.

4.3. Реконструкция поверхности лица и камеры

Для задачи реконструкции трехмерной поверхности лица и нахождения параметров перспективной камеры также были проведены аналогичные эксперименты по исследованию зависимости ошибки от угла поворота лица и сравнению средней ошибки с аналогами. Эксперименты проводились на полностью случайных лицах морфированной модели со всеми случайными коэффициентами $\alpha_i \in N(0, 1), i = 1..199$.

Исследование зависимости от угла проводилось при равномерном повороте одного лица с шагом в 10° в диапазоне от -30° до 30° . Ошибка восстановления параметров морфированной модели $\alpha_i, i = 1..199$, ошибка поворота и переноса представлены на рисунке 8.

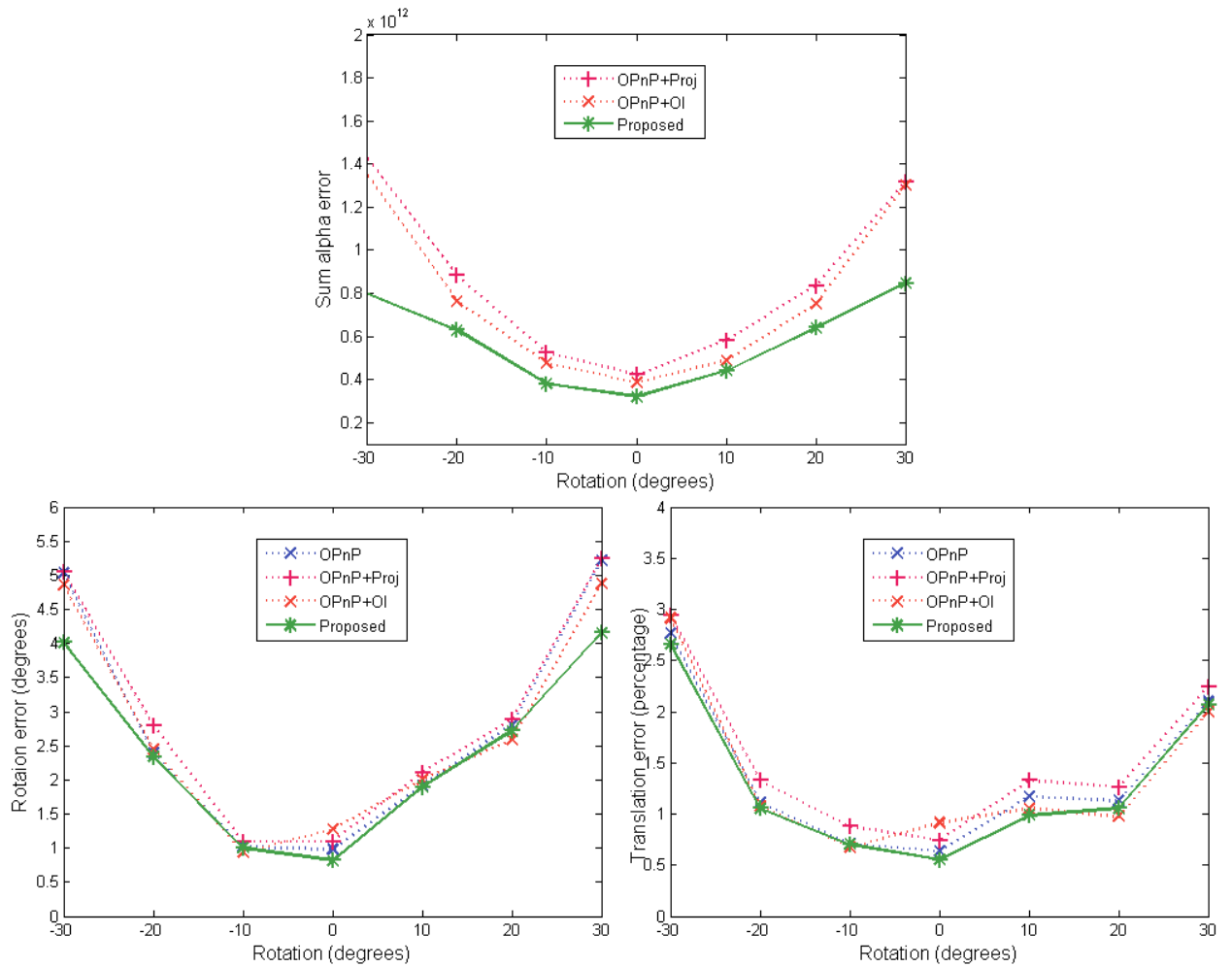


Рис. 8: Зависимость от поворота лица ошибки параметров морфированной модели, поворота и переноса

На графике ошибки поворота и переноса видно, что предлагаемый подход дает ошибку близкую к аналогам, при этом стабильно показывает один из наилучших результатов. Следует отметить, что алгоритм ортогональной ошибки и алгоритм подсчета ошибки перспективного преобразования ухудшают начальные значения поворота найденные OPnP алгоритмом, за счет оптимизации параметров морфированной модели. В отличие от них предлагаемый подход, как и для задачи нахождения камеры только улучшает начальное значение. При этом он точнее аналогов при реконструкции поверхности лица.

Для подсчета среднего и медианы было проведено 50 тестов со случайным поворотом модели лица от -30° до 30° градусов. Ошибки параметров морфированной модели, ошибка матрицы поворота и вектора

переноса представлены на рисунке 9. Медиана ошибок представлена в Приложении 1 на рисунке 13.

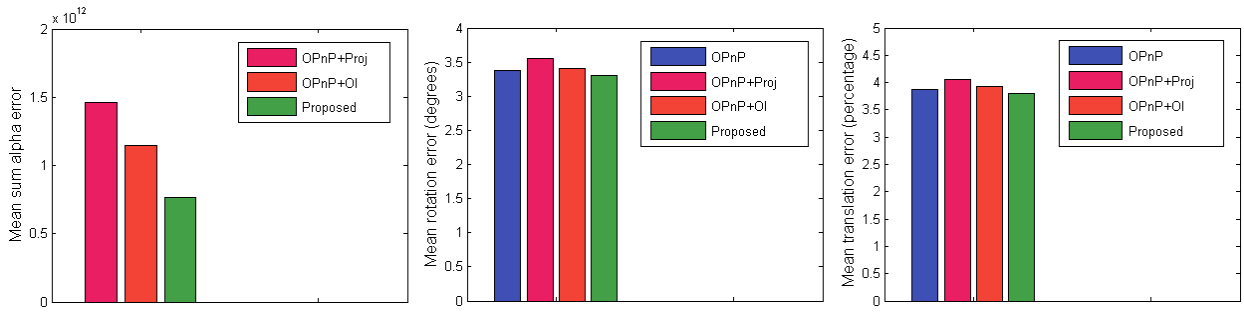


Рис. 9: Средняя ошибка трехмерной поверхности и камеры на 50 тестах

В среднем предлагаемый подход дает наиболее точный результат реконструкции поверхности лица. При этом начальное значение найденное OPnP алгоритмом не ухудшается, а улучшается на небольшое значение. На рисунке 10 показаны примеры реконструированных предлагаемым подходом поверхностей лиц. Слева показано исходное изображение лица, а справа полученная трехмерная модель с учетом найденного поворота и переноса.

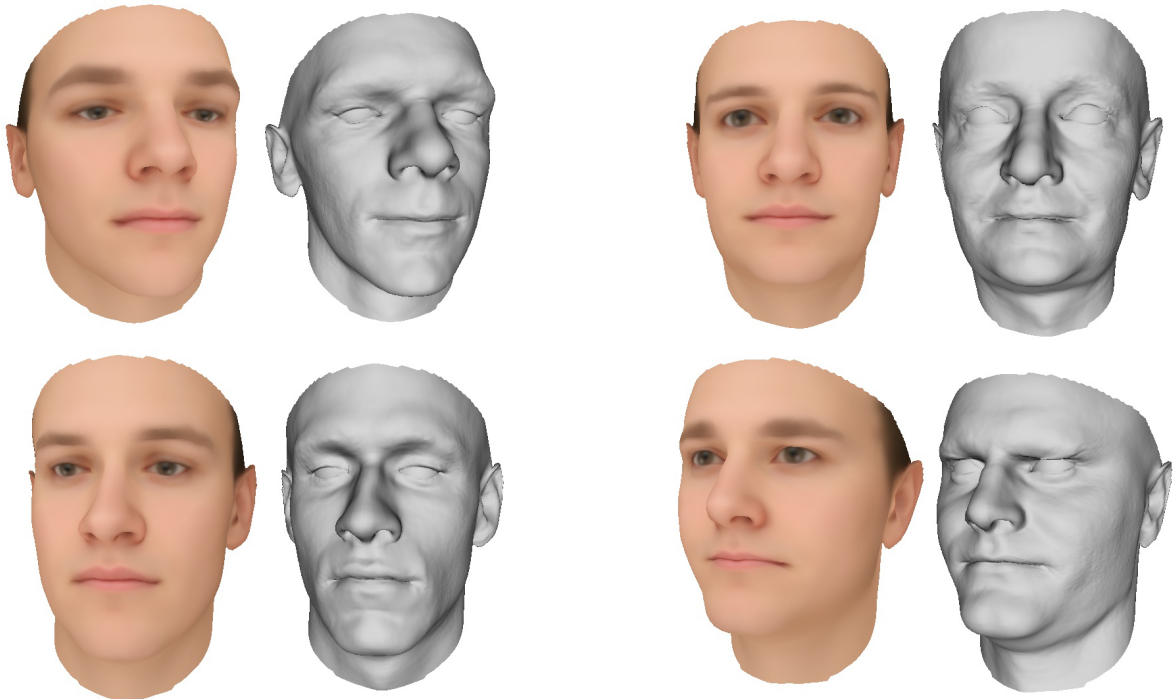


Рис. 10: Реконструированные лица с учетом найденного положения

4.4. Зависимость от параметров тестовых данных

С целью проверки стабильности работы предлагаемого подхода было проведено исследование зависимости результатов от уровня зашумленности исходных изображений. При добавлении случайного гауссовского шума SDM детектор дает смещение ключевой точки в среднем на 4.12 пикселя на зашумленном изображении относительно изображения без шума в среднем на 4.12 пикселя. Исследование проводилось на изображениях разрешения 800 на 600. Исследование ошибки поворота и переноса для задачи нахождения параметров камеры представлено на рисунке 11.

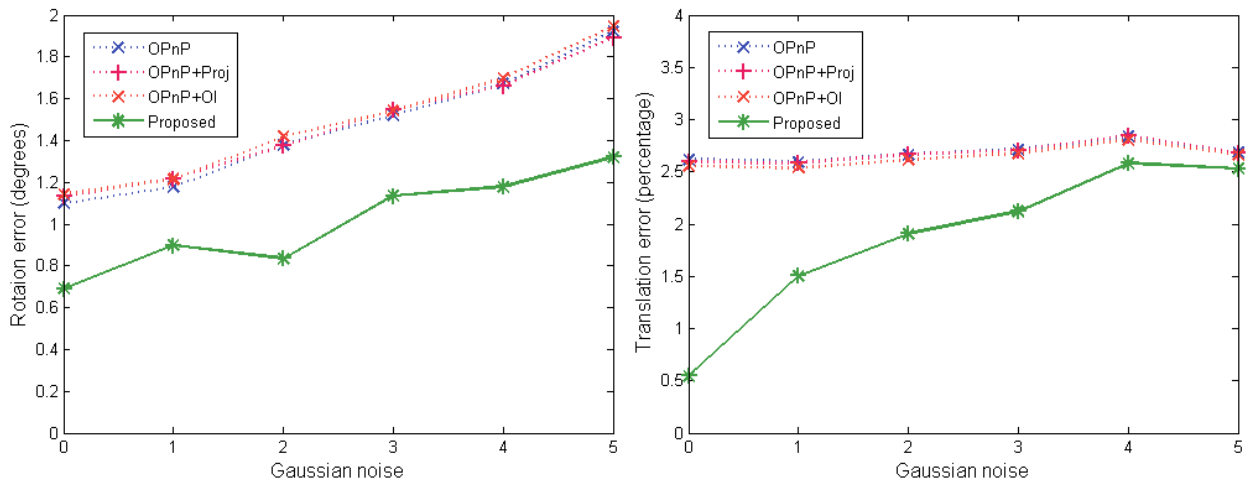


Рис. 11: Ошибка камеры от добавления шума на изображение

Из графиков можно сделать вывод, что ошибка предлагаемого подхода увеличивается с добавлением шума. При этом ошибка поворота аналогов также увеличивается, из-за чего сохраняется разница с предлагаемым подходом. для нахождения переноса аналогично более стабильны. На рисунке 16 в Приложении представлен график зависимости ошибки для задачи реконструкции трехмерной поверхности и камеры.

Заключение

В результате работы были выполнены следующие задачи:

- разработана система сбора статистики распределения ключевых точек в трехмерном пространстве;
- реализован статистический подход к решению задачи реконструкции трехмерного лица по изображению;
- разработана система генерации тестовых данных и проведения экспериментов;
- сравнение показало, что для задачи нахождения камеры подход точнее аналогов в 2 раза, а для задачи реконструкции поверхности лица и камеры дает ошибку по камере близкую к аналогам и в 2 раза меньшую ошибку поверхности лица;
- исследована зависимость результатов предлагаемого подхода от ковариационной матрицы и добавления Гауссовского шума на изображения лиц.

Таким образом, было получено, что предлагаемый статистический подход на базе морфированной модели дает по сравнению с аналогами наиболее стабильные результаты на разных тестовых данных, как для задачи нахождения перспективной камеры, так и для задачи реконструкции поверхности лица и камеры одновременно.

Дальнейшая работа с целью улучшения полученных результатов может быть направлена в сторону применения других алгоритмов детекции ключевых точек, а также увеличения количества используемых в алгоритме ключевых точек за счет добавления неустойчивых особых точек по краям лица и подбородка.

Список литературы

- [1] Blanz Volker, Vetter Thomas. A morphable model for the synthesis of 3D faces // Proceedings of the 26th annual conference on Computer graphics and interactive techniques. / ACM Press/Addison-Wesley Publishing Co. — 1999. — P. 187–194.
- [2] Cootes T. F., et al. Active shape models-their training and application // Computer vision and image understanding. — 1995. — P. 38–59.
- [3] Dalal N., Triggs B. Histograms of oriented gradients for human detection // Computer Vision and Pattern Recognition. — 2005. — P. 886–893.
- [4] DeMenthon D., Davis L. S. Face alignment through subspace constrained mean-shifts // IEEE Transactions on Pattern Analysis & Machine Intelligence. — 1992. — P. 1100–1105.
- [5] Face alignment by explicit shape regression / X. Cao, Y. Wei, F. Wen, J Sun // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. — 2012. — P. 2887–2894.
- [6] Haralick R. M., et al. Pose estimation from corresponding point data // Systems, Man and Cybernetics, IEEE Transactions on. — 1989. — P. 1426–1446.
- [7] Hartley R., Zisserman A. Multiple view geometry in computer vision. — Cambridge university press, 2003.
- [8] Huber P., et al. Fitting 3D Morphable Models using local features // arXiv preprint arXiv:1503.02330. — 2015.
- [9] Izadi S., et al. KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera // Proceedings of the 24th annual ACM symposium on User interface software and technology. / ACM. — 2011. — P. 559–568.

- [10] Jiang D. et al. Efficient 3D reconstruction for face recognition // Pattern Recognition. — 2005. — P. 787–798.
- [11] Jiang D. et al. Park U., Jain A. K. 3D face reconstruction from stereo video // Computer and Robot Vision. The 3rd Canadian Conference on. — 2006. — P. 41–41.
- [12] Kemelmacher-Shlizerman I., Basri R. 3d face reconstruction from a single image using a single reference face shape // Pattern Analysis and Machine Intelligence, IEEE Transactions on. — 2011. — P. 394–405.
- [13] Lepetit V., Moreno-Noguer F., Fua P. Epnp: An accurate $O(n)$ solution to the pnp problem // International journal of computer vision. — 2009. — P. 155–166.
- [14] Li S., Xu C., Xie M. A robust $O(n)$ solution to the perspective-n-point problem // Pattern Analysis and Machine Intelligence, IEEE Transactions on. — 2012. — P. 1444–1450.
- [15] Lienhart R., Maydt J. An extended set of haar-like features for rapid object detection // Image Processing. 2002. Proceedings. 2002 International Conference on. / IEEE. — 2002. — P. 900–903.
- [16] Lu C. P., Hager G. D., Mjolsness E. Fast and globally convergent pose estimation from video images // Pattern Analysis and Machine Intelligence, IEEE Transactions on. — 2000. — P. 610–622.
- [17] Moré J. J. The Levenberg-Marquardt algorithm: implementation and theory. — Numerical analysis : Springer Berlin Heidelberg, 1978.
- [18] Parke F.I. A parametric model for human faces. №. UTEC-CSc-75-047. — UTAH UNIV SALT LAKE CITY DEPT OF COMPUTER SCIENCE, 1974.
- [19] Paysan P. A 3D face model for pose and illumination invariant face recognition // Advanced video and signal based surveillance, 2009.

AVSS'09. Sixth IEEE International Conference on. — 2009. — P. 296–301.

- [20] Qu C., et al. Adaptive Contour Fitting for Pose-Invariant 3D Face Shape Reconstruction. — 2015.
- [21] Saragih J. M., Lucey S., Cohn J. F. Deformable model fitting by regularized landmark mean-shift // International journal of computer vision. — 2011. — P. 200–215.
- [22] Viola P., Jones M. J. Robust real-time face detection // International journal of computer vision. — 2004. — P. 137–154.
- [23] Wang N., et al. Facial feature point detection: A comprehensive survey // arXiv preprint arXiv:1410.1037. — 2014.
- [24] X. Gao, et al. A review of active appearance models // Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on. — 2010. — P. 145–158.
- [25] Xiong X., Torre F. Supervised descent method and its applications to face alignment // Proceedings of the IEEE conference on computer vision and pattern recognition / IEEE. — 2013. — P. 532–539.
- [26] Zheng Y. Revisiting the pnp problem: A fast, general and optimal solution // Proceedings of the IEEE International Conference on Computer Vision. — 2013. — P. 2344–2351.

Приложение

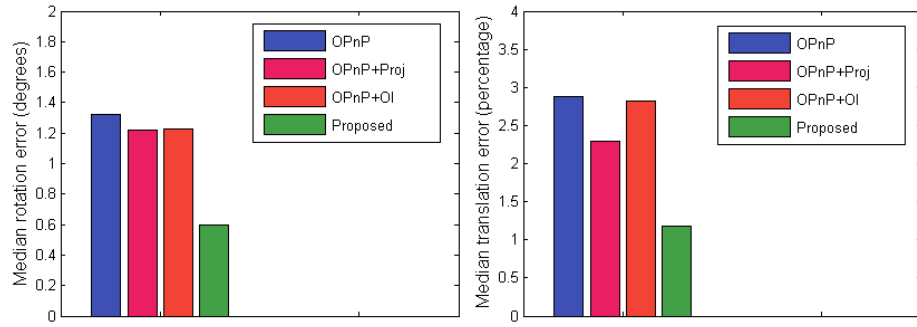


Рис. 12: Медиана ошибки для задачи нахождения камеры

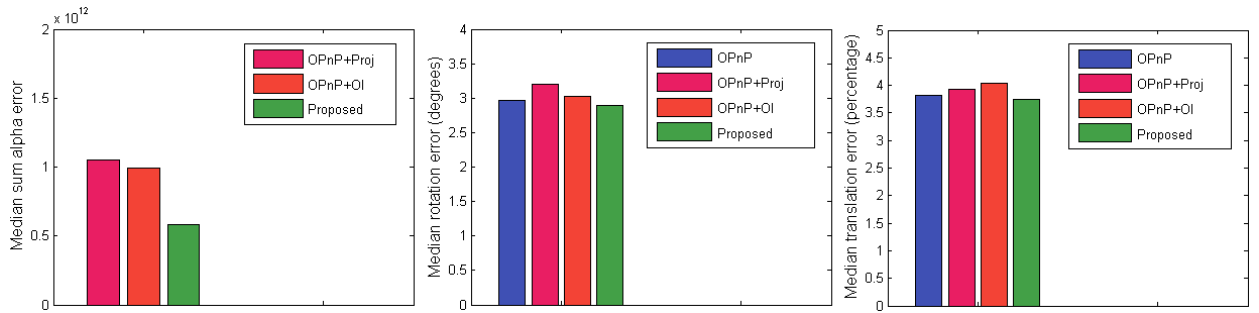


Рис. 13: Медиана ошибки для задачи реконструкции трехмерной поверхности и нахождения камеры

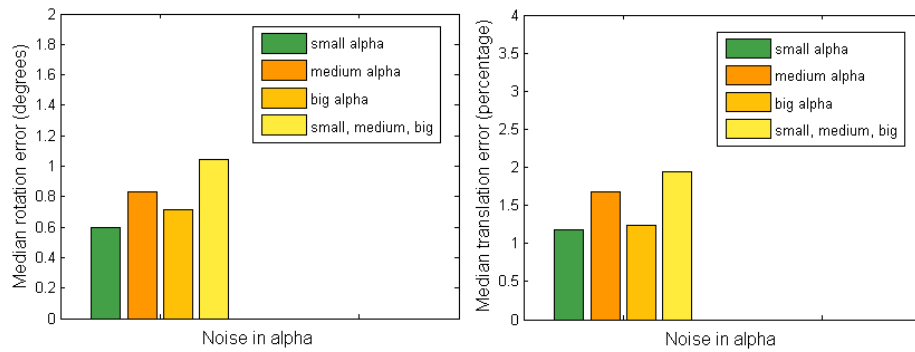


Рис. 14: Медиана ошибки камеры от ковариационной матрицы

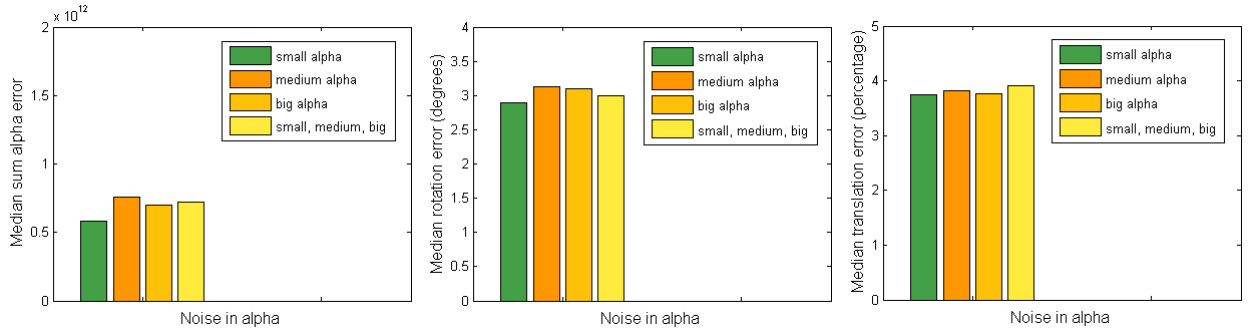


Рис. 15: Медиана ошибки реконструкции поверхности и камеры в зависимости от ковариационной матрицы

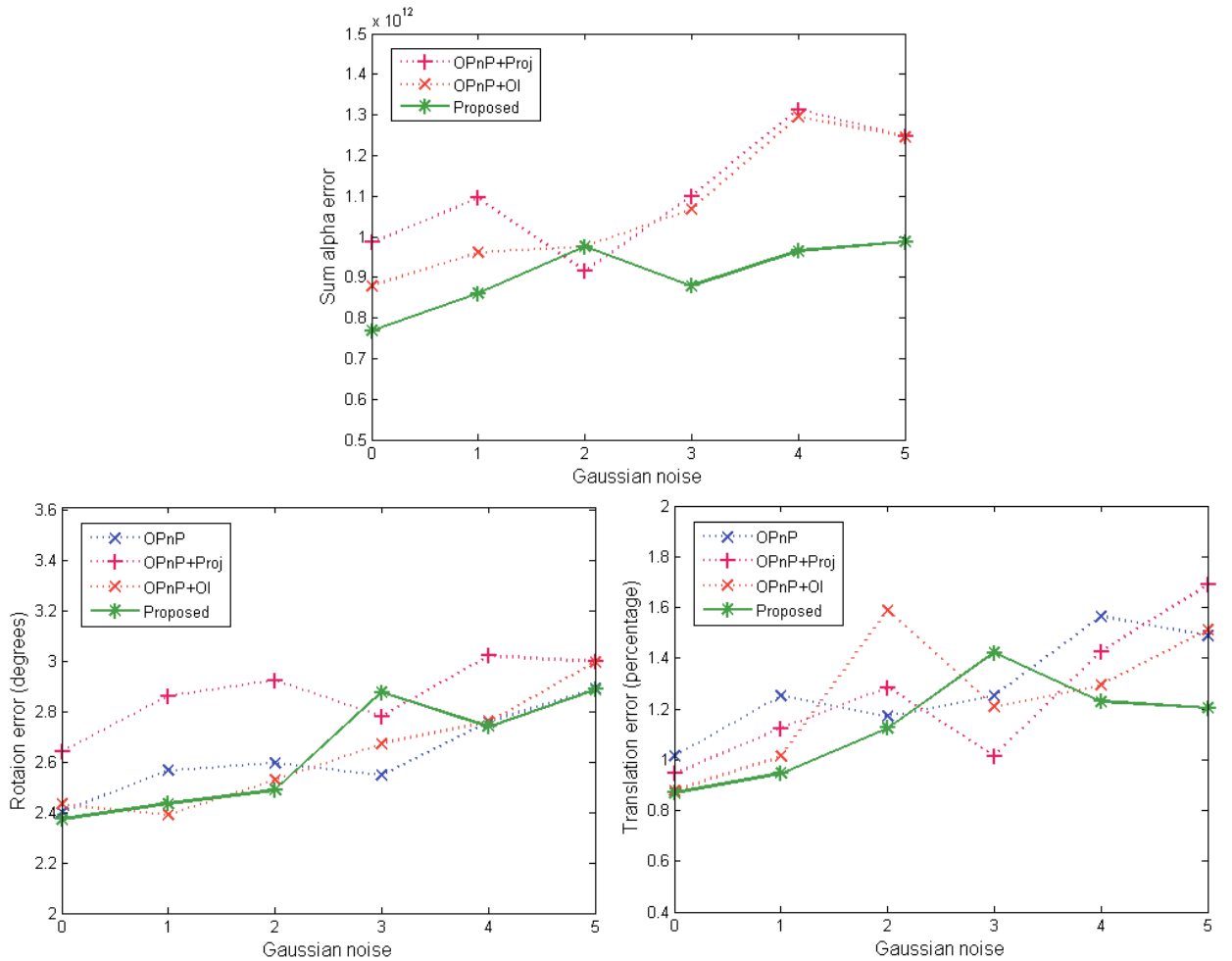


Рис. 16: Ошибка в зависимости от шума на изображениях для задачи реконструкции поверхности и камеры